Analysis of Single Nucleotide Polymorphisms in the Mitochondrial Genome of the Domestic Cat for Forensic Use

By
Cleo Burnett
April, 2013



Allegheny College
Department of Biology

Analysis of Single Nucleotide Polymorphisms in the Mitochondrial Genome of the Domestic Cat for Forensic Use

By
Cleo Burnett
April, 2013

A Senior Comprehensive Project in Partial Fulfillment of the Requirements for a Bachelor of Science Degree from Allegheny College

I hereby recognize and pledge to fulfill my responsibilities as defined in the Honor Code and to maintain the integrity of both myself and the College community as a whole.

_____
Pledge

Approval by the Senior Thesis Committee:

_____
Dr. Kristen Webb

_____
Dr. Lisa Whitenack

Table of Contents

Abstract…… 4

Introduction…….. 4

Research Design and Methods…… 10

Results…….. 12

Discussion……...… 17

Acknowledgements…….. 22

Literature Cited……... 22

Appendix…….. 25

Abstract

Hair is often found at crime scenes and may be used as an identification tool. Due to the strong circular structure and large quantity of mitochondrial DNA, it is often the only usable DNA that can be extracted from hair. While multiple studies have been done on the use of mitochondrial DNA for forensics in dogs, there is relatively little known on this subject in cats. Currently those papers investigating mitochondrial DNA in cats focus on the control region, an approximately 1,560 base pair region which is non-coding and is therefore more free to acquire mutations. A 1,200 bp region of the mitochondrial genome including the *ATP8*, *ATP6*, and *COX3* genes was sequenced from blood samples in an attempt to determine forensic usefulness of loci outside the control region. In a sample size of 12 cats, 10 haplotypes were found and the random match probability was calculated to be 11.1%. These findings show that the region outside the control region could be useful towards identifying the donor of the hair sample and that the cat mitochondrial genome needs to be sequenced in its entirety to accurately assess which areas are most variable.

Introduction

Biological materials such as hair and saliva are often collected as evidence from crime scenes. These materials may be left by humans or animals involved in the crime. Situations where evidence may be directly left behind include occasions where an animal is the victim of a crime (such as dog fighting or animal abuse) or the suspect (a biting incident or having caused a traffic accident). Even if the animal was not present at the crime scene, hairs may be present if they are deposited by someone or something that came in contact with the animal (secondary transfer). When biological evidence is found DNA is often extracted, if

possible, in order to identify who or what it came from. Usable DNA evidence can be successfully collected from as little material as a single hair or a swab of a bite wound (Scharnhorst and Kanthaswamy, 2011). DNA is a powerful forensic tool due to its ability to identify or exclude suspects based on the specific sequence of nucleotides. While differences such as hair color or thickness also vary between individuals, these may vary within the individual and thus do not have the strength of DNA in identifying a source (Houck and Budowle, 2002).

Hair is commonly collected from crime scenes. While the hair shaft is made of dead cells and does not contain much intact nuclear DNA, it contains large quantities of mitochondrial DNA (mtDNA). This is due to the fact that mtDNA is found in hundreds of copies per cell as opposed to the two copies of nuclear DNA. In addition, the circular shape of mtDNA better protects it from degradation compared to the linear nuclear DNA, which helps preserve longer stretches of intact DNA. These facts suggest that even when nuclear DNA is not present in a usable amount, mtDNA may be more readily available and therefore potentially useful as a form of identification or exclusion of an individual as being the source of the DNA (Tarditi et al., 2011, Halverson and Basten, 2005). The use of mtDNA is not limited to hair but the same factors that make it useful when analyzing hair samples make it useful for samples which may have contained much more nuclear DNA originally, but have lost most of this over time due to degradation caused by environmental factors or attempts to destroy the DNA. Despite the relative ease of finding usable mtDNA when compared with nuclear DNA, it is not as strong of a marker for identification of the individual who contributed the DNA. This is in large part due to the direct maternal inheritance of mtDNA and the important functions of the genes to the energy creating process of oxidative

phosphorylation. These factors lead to highly conserved mtDNA sequences which prevent mtDNA from being used to single out a suspected individual whose sequence matches that of a sample as being the definite source of the DNA when it is compared to biological evidence left at the crime scene, though it can be used to rule out those who do not have matching sequences (Schindler et al., 1999).

One particular region of the mtDNA, the control region (mtCR), is often of interest because it is non-coding and therefore has a higher mutation rate than a more conserved coding region, so it is seen as a good locus to use to look for differences in nucleotide sequence. Even with a mutation rate higher than the rest of the mitochondrial genome (mtGenome), different individuals may possess the same sequence, or haplotype, within the mtCR. For example, in a study of 553 dog mtCRs, 143 haplotypes were present. Just over half the dogs shared one of nine common haplotypes with at least 11-59 other individuals in the data set (Webb and Allard, 2009a). Sequencing the highly polymorphic fragment of the mtCR between bases 15,595 and 15,654 in 242 dogs resulted in 23 haplotypes and a random match probability of 18%, indicating that two dogs picked randomly from the population have an 18% chance of having the same haplotype (Baute et al., 2008). Dogs from Sweden, Germany, China, Japan, and the UK, had a random match probability ranging from 5-14% (Angleby and Savolainen, 2005).

Excluding the mtCR, there are approximately 15,458 base pairs of coding region sequence in the dog mtGenome (Kim et al., 1998). While the mtGenome contains some highly conserved loci, it is far larger than the mtCR alone. When looking at a sample of 79 dogs of varying breed and haplotype, 356 base changes (or single nucleotide polymorphisms – SNPs) were found in the mtGenome, with 43% of the SNPs being unique to an individual.

This is comparable to the 42% of SNPs in the mtCR that are unique to an individual. More individuals had unique mtGenome sequences than mtCR sequences, indicating that mtGenome sequence data is more likely to be of use when attempting to identify or rule out a potential source (Webb and Allard, 2009b). Despite this, full intact mtDNA may be difficult to find when samples have not been kept under favorable conditions, preventing initial collection of a complete mtGenome. In addition, sequencing the mtGenome in its entirety requires more time, money, and resource consuming and could be avoided if certain regions were known to be more useful than others for forensic identification.

Despite the fact that as of 2011, there were approximately 8.2 million more cats than dogs owned in the United States according to the Humane Society of the United States (http://www.humanesociety.org/issues/pet_overpopulation/facts/pet_ownership_statistics.html), there have been more studies on the mtCR and mtGenome as a whole in dogs than cats, perhaps due to the fact that dogs may be more likely than cats to cause injury through bites, or be used criminally for fighting, leading to a greater chance to be involved in a criminal case. However, due to the greater population of cats as a whole they may be equally or more likely to have hair left at a crime scene, or have hair left on a suspect by secondary transfer. Most breeds of cat shed hair in large enough quantities that it is difficult to own one or be in the presence of one without a few strands of hair clinging to clothing. In one example of crime in which secondary transfer of cat hair was used as evidence, a woman disappeared from her home in Richmond, Prince Edward Island, Canada. Her car was found abandoned and contained stains of her blood. A men's jacket was also found with her blood on it, as well as some white cat hairs. At the time, the woman's estranged husband had been living with his parents who owned a white cat. Using mtDNA analysis, the cat hair was found to be a match

to the hair in the lining of the men's jacket and the ex-husband was charged with her murder (Menotti-Raymond et al., 1997).

The mtCR in cats is approximately as useful in cats as in dogs in terms of random match probability and formation of haplotypes. A survey of nearly 1400 cats showed 12 major haplotypes representing 83% of the cats, with an additional 8% possessing one of 30 sub-haplotypes derived from a major haplotype. None of the 30 sub-haplotypes were present in large enough numbers to represent more than 0.75% of the total sample size. The remaining 9% of cats had unique haplotypes. However, the haplotype most different from the reference sequence (first sequence published for a particular species) only had 16 SNPs across the entire region, demonstrating how similar this area is between individuals and its limitations as an identification tool; most of the major types differ only by 3 SNPs. The random match probability for the dataset was 11.8%, indicating that the mtCR sequence of an individual has an 11.8% chance of being identical to that of another cat simply by chance (Grahn et al., 2011). Another study using 174 cats found a random match probability of 16.72% (Tarditi et al., 2011). If hair from a crime scene was determined to be one of the more rare haplotypes, it could rule out a large percentage of cats as being the source, and/or people who had recently come into contact with them. The similarities between this data and the data reported for dogs indicate an equal need for further refinement of the process of mtDNA sequencing, leading to lower random match probabilities. While this has progressed to looking at the mtGenome as a whole in dogs, no further research has been done on the potential use of other areas of the mtGenome in cats.

When more of the mtGenome of cats is examined, one would expect to find more SNPs and a greater exclusion capacity, as Webb and Allard (2009b) found in dogs. This

would increase the identification capabilities of evidence in which no or little nuclear DNA is present. The similar random match probability for cats would indicate that SNPs found outside the mtCR may be equally as valuable as previously found in dogs, and may exclude an individual cat from being the source of mtDNA at a crime scene. In most mammals, the mtGenome contains at least13 protein coding, 22 tRNA, and 2 rRNA genes (Lopez et al., 1996). When the mtGenome of other mammals is compared (human, chimp, bonobo, harbor seal, grey seal, blue whale, fin whale, donkey, and horse), patterns in the frequency of mutations over time within certain genes are observed. *ATP8* had the highest mean substitution rate, while *COX1, COX2,* and *COX3* had the lowest. The mean substitution rates for *ATP6, ND6, ND5*, and *ND2* were all higher than the average substitution rate of all the genes studied (Pesole et al., 1999). While this data is viewed through the lens of species evolution, a gene with a higher rate of nucleotide substitution may be a good place to begin the search for a particularly polymorphic area. In the mtGenome of dogs, areas with the highest SNP frequencies outside the mtCR have been found within the *ND4, ND5*, and *COX3* genes. The locus with the greatest percentage of SNPs was at the beginning of the *COX3* gene (Webb and Allard, 2009b). Despite the apparent disagreement with Pesole et al., (1999), the authors mention that a large portion of the *COX3* gene was very highly conserved which would lower the overall substitution rate for the gene and may disguise a shorter but particularly highly variable region. *ATP6* and *ATP8* fall immediately before *COX3* in the feline mtGenome (Lopez et al., 1996), which are both genes that were found to have a high substitution rate over time (Pesole et al., 1999). If the *ATP8, ATP6,* and *COX3* genes in the cat mtGenome contain SNPs that form many haplotypes, it may suggest that the entire

mtGenome be sequenced and analyzed to look for other mutational hotspots as Webb and Allard (2009b) have done in dogs.

Research Design and Methods

Sample Collection

Cat blood samples were donated by Conneaut Lake Veterinary Hospital (Conneaut Lake, PA). Information including the breed, age, and sex was collected for each donor. Cats used were not maternally related as the mitochondrial genome is passed directly from mother to offspring. All samples were frozen at -20°C until extraction.

DNA Extraction

DNA was extracted from 100ul of blood the DNeasy Blood and Tissue DNA Extraction Kit (Qiagen, Germantown, MD), following the instructions for Purification of Total DNA from Animal Blood or Cells (Spin-Column Protocol) and including the optional RNAse treatment. A negative control was included to show if contamination had occurred. Success of each DNA extraction was confirmed by checking for presence of DNA after 1% agarose gel electrophoresis. Concentration and purity of the extracted DNA was determined using Nanodrop spectroscopy.

Primers and Polymerase Chain Reaction (PCR)

Primers were designed (Table 1) to amplify 1,400 bases including part of the *ATP8* gene, the entire *ATP6* gene and the start of the *COX3* gene (mtGenome bases 8,700 – 10,100), using the only domestic cat mtDNA genome sequence posted on Genbank (accession number NC_001700). Primer3 (Rozen and Skaletsky, 2000) was used to generate primers, and

complementarity and binding sites to the mtGenome were confirmed through pairwise alignment to the mtGenome using BLAST (www.ncbi.nlm.nih.gov/blast).

Table 1. Primer information for primers used in PCR and sequencing.

Primer Name	Primer Sequence (5' – 3')	5' Coordinate	3' Coordinate
Fd_mt_F	TGCCACAACTAGATACATCCAC	8659	8680
Fd_mt_R	ATCCTGTGGCCATGAAGAAG	10,105	10,086

Primers were used to set up a 50ul polymerase chain reaction (PCR) for each sample using 100ng of DNA, 25ul of Quick-Load Taq 2x master mix (New England Biolabs, Ipswich, MA), and 1.26uM/ul of each primer. Volume was adjusted to 50ul with water as needed. A negative control replacing DNA with water was used to check for contamination. PCR reactions were placed in the thermal cycler programmed for 5 minutes at 96°C, followed by 40 cycles of 94°C for 45 sec/52°C for 30 sec/72°C for 90 sec. After the cycles, the samples will be held at 72°C for 5 minutes before lowering to a holding temperature of 4°C. PCR success, as well as the size and concentration of each product, was tested by checking for the presence of a single bright band after electrophoresis in a 1% agarose gel including a low mass ladder (Invitrogen, Carlsbad, CA). Successful PCR products were purified using ExoSAP-IT (Affymetrix, Inc., Santa Clara, CA) to remove excess primers and dNTPs, following the manufacturer's protocol, prior to DNA sequencing.

Sequencing and Sequence Editing

Successful PCR products were sent to Eurofins MWG Operon (Huntsville, AL) for sequencing in a company provided barcoded sample tube, along with an adequate amount of primer following suggestions given by the company (http://www.operon.com/services/dna-

sequencing/sample-prep.aspx). Sequencher (v5.0.0, Gene Codes Corporation, Ann Arbor, MI) was used to view and edit the resultant sequences.

Identification of SNPs

Each newly collected sequence was aligned to the reference sequence (NC_001700) using Sequencher v5.0.0, according to the standards set by Wilson et al., 2002. These alignments identified the position and base change of all SNPs in present in the newly collected sequences relative to the reference sequence. Results were recorded in an Excel spreadsheet. This data was used to determine haplotypes and the random match probability (Σx_i^2 , where X_i is the frequency of the ith haplotype), as well as SNP frequencies per gene.

Results

Sample Collection

Twenty-two cat blood samples were collected from Conneaut Lake Veterinary Hospital from unrelated cats (Table 2).

Table 2. Cat blood samples. [a]

Sample Name	Breed	Sex	Age (years)
FC1	DSH	Male (neutered)	2
FC2	DSH	Male (neutered)	16
FC3	DSH	Female (spayed)	16
FC4	DSH	Male	6
FC5	DSH	Male (neutered)	14
FC6	DSH	Male (neutered)	16
FC7	DSH	Female (spayed)	15
FC8	DSH	Male	5
FC9	DSH	Female (spayed)	17
FC10	DSH	Male (neutered)	7
FC11	Siamese mix	Female (spayed)	4
FC12	DSH	Female (spayed)	4
FC13	DSH	Male (neutered)	8
FC14	DSH	Female (spayed)	15
FC15	DSH	Female (spayed)	8

FC16	DSH	Male (neutered)	4
FC17	DLH	Male (neutered)	1
FC18	DLH	Male (neutered)	9
FC19	DSH	Male (neutered)	5
FC20	DSH	Male	0.8
FC21	DLH	Male (neutered)	15
FC22	DLH	Male (neutered)	10

[a] Greyed rows indicate that the sample was not used in the final analysis due to a failure to extract or amplify mtDNA.

DNA Extraction

DNA extraction initially failed due to contamination with ethanol during the wash steps of the extraction protocol. When repeated properly, 21 out of 22 samples extracted with great enough quality and quantity to perform PCR (Table 3). Any sample with a concentration greater than 7ng/ul was used for PCR.

Table 3. Results of DNA extraction. [b]

Sample	Concentration (ng/ul)	Purity (260/280)	Sample	Concentration (ng/ul)	Purity (26/280)
FC1	14.9	1.96	**FC12**	68.8	1.83
FC2	12.2	1.80	**FC13**	24.4	1.79
FC3	38.6	1.84	**FC14**	24.4	1.81
FC4	22.7	1.84	**FC15**	13.0	1.62
FC5	11.0	1.79	**FC16**	25.7	1.81
FC6	11.4	1.90	**FC17**	21.8	1.80
FC7	3.2	1.40	**FC18**	21.0	1.84
FC8	9.7	1.62	**FC19**	10.6	1.71
FC9	21.7	1.89	**FC20**	29.0	1.74
FC10	9.1	1.89	**FC21**	8.5	1.89
FC11	10.9	1.61	**FC22**	28.2	2.00
control	1.4	0.92	**control**	2.2	6.53

[b] Greyed rows indicate that the sample was not carried on to PCR due to a failure to extract.

Primers and Polymerase Chain Reaction (PCR)

All 21 samples that extracted successfully were used in PCR. Of these, 12 produced a bright enough band on the gel to suggest the presence of enough DNA for sequencing (these samples are not greyed out in Table 2). When compared to the low mass ladder all visible

bands were of expected size (~1,400 bps) and had concentrations around 80ng. Lanes with no bands, or bands that appeared to have a concentration lower than 60ng were not sent for sequencing. In general all visible bands had a concentration near 80ng, with no or few bands under this.

Sequencing and Sequence Editing

Samples sequenced with the primer Fd_mt_F returned sequence data with very short peaks on the chromatogram. Samples sequenced with Fd_mt_R were sent with twice the amount of primer to attempt to troubleshoot the chromatography issues, which were suspected to be an issue related to the primer to DNA ratio, but the returned data was of very short length. No further troubleshooting was performed due to time constraints. While sequence pairs did not cover the entire desired 1,400 base pair region, the data was still of use and ranged from to 251-992 bases in length with at least half of the sequences having a length of 688 or more. Most sequence pairs covered at least 86.7% of the target region per individual with a gap of varying length at the end of *ATP6*. Minimal editing was required (for sequences see Appendix A).

Identification of SNPs

In total there were 35 loci with at least 1 SNP present (Table 4). 154 SNPs were found across these 35 loci in all 12 samples. *ATP8* contained 10 of the total 35 loci, while *ATP6* had 9 and *COX3* had 16 (Figures 1-3). The 12 samples formed 10 haplotypes with a random match probability of 11.1%. *ATP8, ATP6,* and *COX3* had SNP frequencies of 1 in 50.4, 1 in 114.9, and 1 in 132.0, with an overall average of 1 SNP in every 105.9 bases over the entire sequenced region.

Table 4. SNPs in the newly collected samples relative to the reference sequence.

	8714	8726	8727	8731	8734	8735	8737	8759	8800	8801	8853	8902
Ref	A	A	T	A	T	A	A	A	G	G	C	C
FC1	.	.	.	.	.	.	.	.	T	T	.	T
FC3	C	T	G	C	G	.	T	.	T	T	A	T
FC4	.	.	.	.	.	.	.	.	T	T	.	T
FC12	.	.	.	.	.	.	.	.	T	T	.	T
FC13	.	.	.	.	.	.	.	.	T	T	.	T
FC15	.	.	.	.	.	.	.	G	T	T	A	T
FC16	.	.	.	.	.	.	.	.	T	T	.	T
FC17	.	.	G	.	.	.	.	.	T	T	.	T
FC18	.	.	.	.	.	.	.	.	T	T	.	T
FC19	.	.	.	.	.	G	.	.	T	T	.	T
FC20	.	.	.	C	.	.	T	.	T	T	.	T
FC22	.	.	.	.	.	.	.	.	T	T	.	T

	8906	8972	9110	9211	9277	9346	9367	9573	9614	9634	9647
Ref	T	C	T	A	T	G	C	C	T	G	A
FC1	.	T	C	.	C	A	.	.	C	.	.
FC3	.	T	C	G	C	A	.	.	C	T	.
FC4	.	T	C	.	C	A	.	.	C	.	C
FC12	.	T	C	G	C	A	.	.	C	.	.
FC13	.	T	C	G	C	A	.	T	C	.	.
FC15	.	T	C	.	C	A	.	.	.	.	.
FC16	.	T	C	G	C	A	.	.	C	.	.
FC17	.	T	C	.	C	A	.	.	C	.	.
FC18	.	T	C	G	C	A	.	.	C	.	.
FC19	C	T	.	.	C	A	T	.	C	.	.
FC20	.	T	C	G	C	A	.	.	C	.	.
FC22	.	T	C	.	C	A	.	.	C	.	.

	9649	9694	9699	9736	9780	9814	9852	9957	9978	10008	10026	10055
Ref	T	A	C	C	C	C	A	G	T	G	C	T
FC1	C	.	.	.	.	.	.	.	.	.	T	.
FC3	C	C	T	T	T	.	.	.	.	A	T	.
FC4	.	.	.	.	.	T	.	.	.	.	T	.
FC12	C	.	.	.	.	.	.	.	.	A	T	.
FC13	C	.	.	.	.	.	.	.	.	A	T	.
FC15	.	.	.	.	.	.	.	.	.	.	T	C
FC16	C	.	.	.	.	.	.	.	.	A	T	.
FC17	C	.	.	.	.	.	.	.	.	.	T	.
FC18	C	.	.	.	.	.	.	.	.	A	.	.
FC19	C	.	.	.	.	.	G	A	C	.	.	.
FC20	C	.	.	.	.	.	.	.	.	A	T	.
FC22	C	.	.	.	.	.	.	.	.	.	T	C

Reference data above collected sample data. Base position of SNP listed in top row, sample name in leftmost column. Dots indicate a match to the reference sequence. Genes separated by bold vertical line in order of *ATP8, ATP6, COX3*.

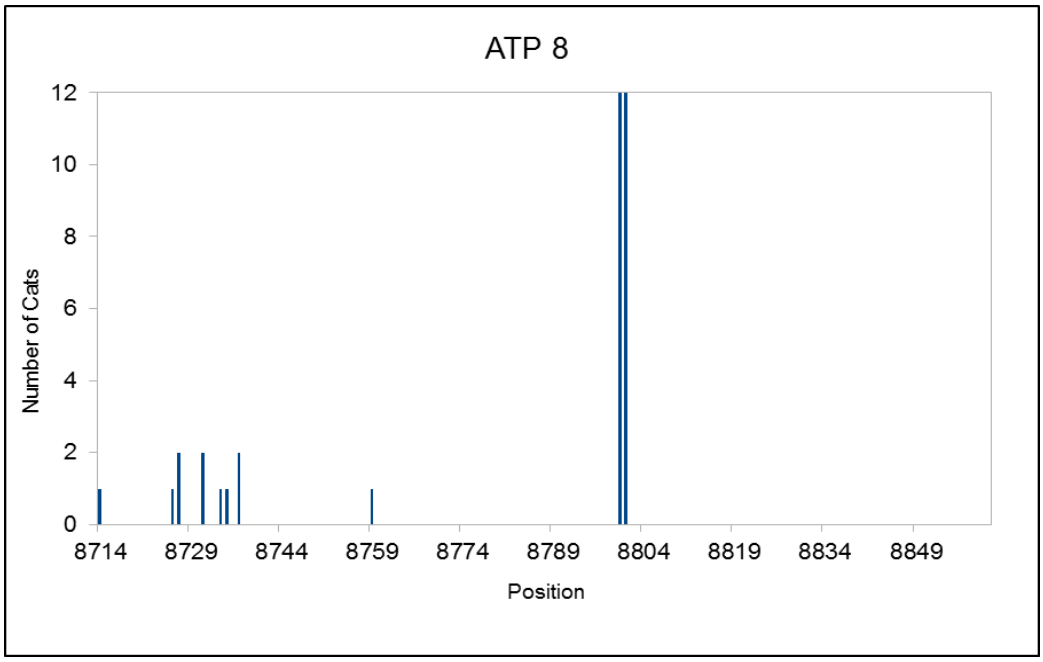

Figure 1. Graph indicating number of cats sharing a SNP at each locus across the sequenced region of *ATP8*.

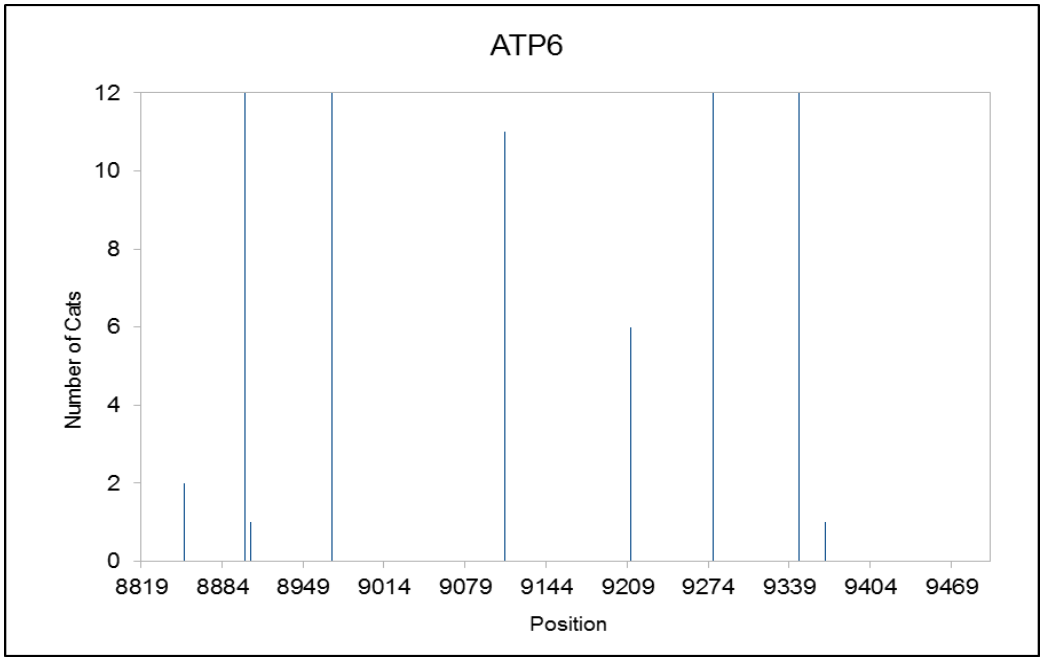

Figure 2. Graph indicating number of cats sharing a SNP at each locus across the sequenced region of *ATP6*.

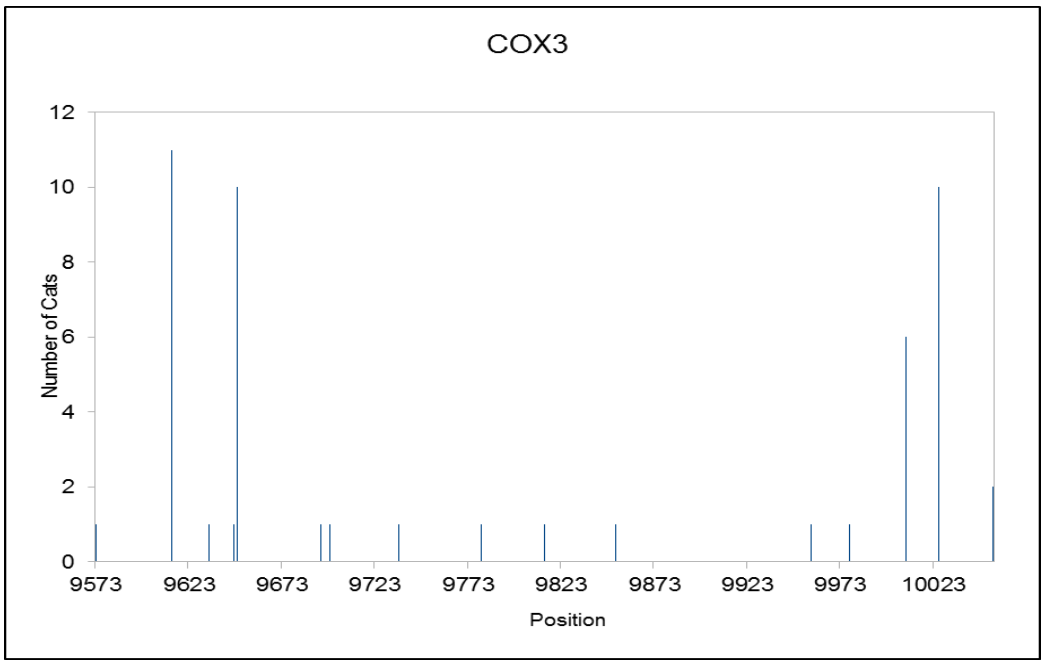

Figure 3. Graph indicating number of cats sharing a SNP at each locus across the sequenced region of *COX3*.

Discussion

The purpose of this study was to look for a potential mutational hotspot within the mtGenome of the domestic cat to be used alongside the mtCR in forensic identification of a sample in which nuclear DNA cannot be used due to degradation. While the region chosen for amplification in this study contains the nucleotides of three protein coding genes, and is therefore expected to be more conserved than the non-coding mtCR, the homologous region was found to contain a large number of SNPs in the domestic dog mitochondrial genome (Webb and Allard, 2009b). SNPs in coding regions are less likely due to the potential for non-synonymous, potentially deleterious mutations. Because of the increased selective pressure to maintain the coding-region sequence, the frequencies of SNPs in these regions are lower relative to non-coding regions, meaning that SNPs in these regions are more likely to

be unique to an individual or small group of individuals. This fact makes these rare SNPs more useful when attempting to rule out an individual as a potential source of biological evidence.

At first, the average frequency of SNPs per gene seems to tell very little about the individual. While SNP frequency information for the feline mtCR is not provided by either group which has studied this region (Grahn et al., 2011, Tarditi et al., 2011), the average in dogs has been found to be 1 in 15 bases (Webb and Allard, 2009a). In comparison the SNP rate for the canine mtGenome aside from the mtCR in was 1 in 50 (Webb and Allard, 2009b). Either the region of the feline mtGenome sequenced here has less SNPs than average, possibly due to the rate for cat mtDNA mutations being under different selective pressures and less than in dogs. Stronger purifying selection may occur in cats due to lifestyle differences between canine pack life and the more individual predatory life of feline ancestors. An individual weakened by a deleterious mutation affecting overall levels of energy may have a higher chance at survival if supported by a group of healthy individuals, such as a pack. Wild cats do not have the safety net of the group to fall back on, and so individual fitness may become more important. If the cat mtGenome has less variation between individuals than has been observed in dogs, it may be a weaker tool to use in forensic investigations where cat mtDNA is present as it may raise the random match probability. This would mean that individual cats would be more likely to have the same haplotype. However, having a smaller amount of variation would be useful when the sample at a crime scene is of a rarer haplotype. In a case such as this, because the haplotype is rare to begin with, there will be fewer cats that the evidence could have been left by.

In addition to the number of SNPs per gene, the location of each SNP within the gene is also worthy of consideration. While the SNPs were roughly evenly distributed throughout the *ATP6* and *COX3* genes there was an apparent mutational hotspot near the beginning of the sequenced region of *ATP8* (Figure 1). In a 45 base region spanning positions 8,714 to 8,759 there were 10 positions where at least one cat possessed a SNP producing a SNP frequency of 1 in 45 bases. Despite this area appearing to be a hotspot, its SNP frequency only slightly surpasses that of the total sequenced area of that gene (1 in 50). If only the 45 base pair region is examined, the 12 cats form 5 haplotypes with a random match probability of 47.2%. If this region is excluded, the cats form 9 haplotypes with a random match probability of 13.9%. Overall, there were no particularly polymorphic regions that could be termed a mutational hotspot, including the region of *COX3* that was found to have the third highest SNP frequency in dogs (Webb and Allard, 2009b).

The 12 cats in the data set formed 10 haplotypes. This means most cats were singletons, with only two groups containing 2 cats each. The random match probability was calculated to be 11.1%. In comparison, the random match probability for regions of the mtCR in cats has been found to be 16.72% in a sample size of 174 (Tarditi et al., 2011) and 11.9% in another study comparing 1394 cats (Grahn et al., 2011). Although it would appear that the region sequenced here has a random match probability equal to or less than the mtCR, this may be an error due to the small sample size. Sequencing more cats would lead to the discovery of more haplotypes which would lower the random match probability, but depending on sample size and number of new haplotypes formed it could also raise the random match probability.

The forensic usefulness of dog mtCR sequences has already been shown and is an acceptable form of evidence (Schneider et al., 1999, Scharnhorst and Kanthaswamy, 2011, Halverson and Basten, 2005). Although cat mtDNA has also been used forensically (Menotti-Raymond et al., 1997, Halverson and Basten, 2005) it is clear that the greater pool of dog knowledge currently makes canine evidence a stronger method of exclusion than feline, emphasizing the need for more a more extensive understanding of feline mitochondrial genetics. While currently published data on cats cannot match the 4.3% random match probability found by Webb and Allard (2009b) using ~1,000 bp of the mtCR in dogs, it is comparable to studies focusing on shorter regions of 60-608 bp which found a random match probability ranging from 10-18% (Baute et al., 2008, Himmelberger et al., 2008, Salvolainen et al., 1997, Angleby and Salvolainen, 2005). These studies of shorter length may be a more meaningful method of comparison between the current data on cats and dogs as they more closely match in terms of the lengths of sequences investigated. A comparison between similar studies suggests the mtCR in dogs and cats appears to be of equal use when calculating the random match probability; however this statement cannot be made accurately without a feline dataset similar in terms of both sample size and mtCR coverage to that of Webb and Allard (2009a).

Webb and Allard (2009b) found that the mtGenome data helped to break up the larger mtCR haplotypes into smaller ones. The creation of more haplotypes lowers the random match probability and thus increases the strength of any biological evidence analyzed in this fashion. Many of the more common feline mtCR haplotypes have a frequency between 15-30% in multiple populations defined by region. Half of the cats from Iran/Iraq grouped within a single haplotype (Grahn et al., 2011). These high haplotype frequencies reduce the

usefulness of mtDNA and require a stronger method of exclusion to be of use forensically. In dogs, after the area outside the mtCR was considered, the largest mtCR haplotype was divided in half (Webb and Allard, 2009b). Potentially doubling the exclusion ability of a particular haplotype is a large improvement and would be very useful in regions where a single haplotype constitutes 50% of the population. The goal of further studies in the mtGenome of cats would be to reproduce this ability for exclusion in cats. Ideally, this requires full mtGenome sequencing for each cat in the data set for areas both inside and outside the mtCR in order to compare the effects of specific SNPs on haplotype division.

Another factor to consider is the quantity of mtDNA available from a biological source of evidence. Blood or tissue samples are best, but it may be that only a single hair was found at a crime scene. Less than 10ng of mtDNA may be extracted from a single shed hair (Vigilant, 1999), preventing the entire mtGenome from being sequenced. In cases such as this, it would be of value to know the regions of mtDNA most useful in haplotype identification. The mtCR is currently used, but could also use a greater resolution of haplotypes as previously mentioned. This further demonstrates the importance of finding the most useful regions for haplotype determination in both cats and dogs.

While all aspects of this study would benefit from an increased sample size, it would also be of interest to sequence additional regions of the mtGenome in these cats. In particular, the mtCR of the cats used in this study should be looked at to compare the usefulness of these regions to the mtCR and see if they are truly of use in identifying a sample or splitting either of the haplotypes that contained more than one individual. On a larger scale the complete mtGenome, including the mtCR, should also be sequenced and the presence, location, and usefulness of particular SNPs in determining haplotypes evaluated.

Acknowledgements

I would like to thank the Conneaut Lake Veterinary Hospital for their donation of blood samples, as well as the Allegheny College Biology Department for funding this project. I would also like to thank Dr. Webb for her assistance with both lab work and the writing of this paper, and my comp lab group for assisting in troubleshooting.

Literature Cited

Angleby H, Savolainen P. 2005. Forensic informativity of domestic dog mtDNA control region sequences. Forensic Sci Int 154(2–3):99-110.

Baute DT, Satkoski JA, Spear TF, Smith DG, Dayton MR, Malladi VS, Goyal V, Kou A, Kinaga JL, Kanthaswamy S. 2008. Analysis of forensic SNPs in the canine mtDNA HV1 mutational hotspot region. J Forensic Sci 53(6):1325-33.

Grahn RA, Kurushima JD, Billings NC, Grahn JC, Halverson JL, Hammer E, Ho CK, Kun TJ, Levy JK, Lipinski MJ, et al., 2011. Feline non-repetitive mitochondrial DNA control region database for forensic evidence. Forensic Sci Int -Genet 5(1):33-42.

Halverson JL, Basten C. 2005. Forensic DNA identification of animal-derived trace evidence: tools for linking victims and suspects. Croatian Medical Journal 46(6):598 605.

Himmelberger AL, Spear TF, Satkoski JA, George DA, Garnica WT, Malladi VS, Smith DG,

Webb KM, Allard MW, Kanthaswamy S. 2008. Forensic utility of the mitochondrial hypervariable region 1 of domestic dogs, in conjunction with breed and geographic information. J Forensic Sci 53(1):81-9.

Houck MM, Budowle B. 2002. Correlation of microscopic and mitochondrial DNA hair comparisons. J Forensic Sci 47(5):964-967.

Kim KS, Lee SE, Jeong HW, Ha JH. 1998. The complete nucleotide sequence of the domestic dog (Canis familiaris) mitochondrial genome. Mol Phylogenet Evol 10(2):210–20.

Lopez JV, Cevario S, O'Brien SJ. 1996. Complete nucleotide sequences of the domestic cat (Felis catus) mitochondrial genome and a transposed mtDNA tandem repeat (Numt) in the nuclear genome. Genomics 33(2):229-46

Rozen S, Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist Programmers. Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386

Salvolainen P, Rosen B, Holmberg A, Leitner T, Uhlen M, Lundeberg J. 1997. Sequence analysis of domestic dog mitochondrial DNA for forensic use. J Forensic Sci 42(4):593-600.

Schnieder PM, Seo Y, Rittner C. 1999. Forensic mtDNA hair analysis excludes a dog from having caused a traffic accident. International Journal of Legal Medicine 112(5):315 16.

Tarditi CR, Grahn RA, Evans JJ, Kurushima JD, Lyons LA. 2011. Mitochondrial DNA sequencing of cat hair: An informative forensic tool. J Forensic Sci 56:S36-46.

Vigilant L. 1999. An evaluation of techniques for the extraction and amplification of DNA from naturally shed hairs. Biol Chem 380:1329-1331.

Webb KM, Allard MW. 2009a. Identification of forensically informative SNPs in the domestic dog mitochondrial control region. J Forensic Sci 54(2):289-304.

Webb KM, Allard MW. 2009b. Mitochondrial genome DNA analysis of the domestic dog: identifying informative SNPs outside of the control region. J Forensic Sci 54(2):275-88.

Wilson MR, Allard MW, Monson K, Miller KW, Budowle B. 2002. Recommendations for consistent treatment of length variants in the human mitochondrial DNA control region. Forensic Sci Int 129(1):35–42

Appendix

Appendix A. Collected cat mtDNA sequences in FASTA format.

>FC1_Fd_mt_F

ATCTCAAAATACTTATATCCATCAAACCCAGAACCTAAATCCATAACCAC

ACTAAAACAACTTAATCCCTGAGAAAAAAAATGAACGAAAATCTATTCGC

CTCTTTCACTACCCCAACAATAATAGGATTACCTATTGTTATTTTAATTA

TTATATTTCCAAGTATTTTATTCCCTTCACCTAACCGACTAATTAATAAC

CGTCTAGTTTCACTCCAACAATGACTAGTACAATTAACATCAAAACAAAT

ACTGGCTATTCATAATCATAAAGGACAAACCTGAGCCCTAATACTAATGT

CCCTAATTCTATTTATTGGGTCAACAAACTTATTAGGCCTATTACCCCAC

TCATTCACCCCAACTACCCAACTATCAATAAATTTAGGAATAGCTATCCC

GCTATGAGCCGGCACTGTAATTACCGGGTTTCGCCACAAGACTAAAGCAT

CTCTGGCCCACTTTCTACCACAAGGAACACCTGTCCCCCTAATTCCTATG

CTTGTAGTCATTGAGACTATCAGCCTCTTTATCCAACCCATAGCTCTCGC

CGTACGACTTACAGCCAACATCACCGCAGGTCACTTATTAATACATCTAA

TTGGAGGAGCCGCCCTAGCCCTGATAAACATTAGCACCTCTATTGCCTTA

ATCACCTTTACCATTCTCATTTTACTAACAATCCTTGAATTTGCCGTAGC

CCTAATCCAAGCCTATGTTTTTACCCTGCTAGTAAGCCTATACTTACATG

ATAACACCTAA

>FC1_Fd_mt_R

GAAGGAGTGTAAAGTAGACCCCTAAAGAGATTGTAATAAATAGTGCTTGA

AGCATGTGTTTTCGATTTCCCTCCATCAAACTGTGGTGAGCTCAGGTGAT

TGATACTCCGGAGGCTAGAAGCACGGAGGTATTAAGTAGTGGAACTTCCA

GGGGGTTCAGGGGAATAATGCCTGTTGGTGGTCAGCATCCTCCTAGCTCG

GGGGTTGGGGCTAGGCTTGAGTGGTAGAAGGCCCAGAAGAAGCCTGCGAA

AAAGAATACTTCTGAGATGATAAAGAGGATTATTCCGTATCGAAGGCCTT

TTTGAACGATAGGTGTATGATGGCCTTGGAATGTGCTTTCTCGGATAATG

TCTCGTCATCATTGATATATAGTTAGTAGGTTGGTAGTTATTCCAAGGGT

TAACAGCAGTGTTGAGTTGTAGTGGAATCATATAGCCAGGCCTGAG

>FC3_Fd_mt_F

CCTATTTATTGTTGTCCCACGAATAATCTCAAAATACTTATATCCATCAA

ACCCAGAACCTAAATCCATAACCACACTAAAACAACTTAATCCCTGAGAA

AAAAAATGAACGAAAATCTATTCGCCTCTTTCACTACCCAAACAATAATA

GGATTACCTATTGTTATTTTAATTATTATATTTCCAAGTATTTTATTCCC

TTCACCTAACCGACTAATTAATAACCGTCTAGTTTCACTCCAACAATGAC

TAGTACAATTAACATCAAAACAAATACTGGCTATTCATAATCATAAAGGA

CAAACCTGAGCCCTAATACTAATGTCCCTAATTCTATTTATTGGGTCAAC

AAACTTATTAGGCCTATTACCCCACTCATTCACCCCAACTACCCAACTAT

CAATAAATTTAGGAATAGCTATCCCGCTATGAGCCGGCACTGTAATTACC

GGGTTTCGCCACAAGACTAAAGCATCTCTGGCCCACTTTCTACCACAGGG

AACACCTGTCCCCCTAATTCCTATGCTTGTAGTCATTGAGACTATCAGCC

TCTTTATCCAACCCATAGCTCTCGCCGTACGACTTACAGCCAACATCACC

GCAGGTCACTTATTAATACATCTAATTGGAGGAGCCGCCCTAGCCCTGAT

AAACATTAGCACCTCTATTGCCTTAATCACCTTTACCATTCTCATTTTAC

TAACAATCCTTGAATTTGCCGTAGCCCTAATCCAAGCCTATGTTTTTACC

CTGCTAGTAAGCCTATACTTACATGATAACACCTAATGACCCACCAAACC

CATGCATACCACATAGTCAACCCTAGCCCATGGCCACTTACAGGAGCCCT

TTCAGCCCTCTTAATAACCTCAGGCCTGGCTATATGATTCCACTACAACT

CAACACTGCTGTTAACCCTTTGAATAACTACCAACCTACTAACTATATAT

CAATGATGACGAGACATTATCCGAGAAAGCCCATTTCAAG

>FC3_Fd_mt_R

GAAGGAGTGTAAAGTAGACTCCTAAAGAGATTGTAATAAATAGTGCTTGA

AGCATGTGTTTTCGATTTCCCTCCATCAAACTGTGGTGAGCTCAGGTGAT

TGATACTCCGGAGGCTAGAAGCACGGAGGTATTAAGTAGTGGAACTTCCA

GGGGGTTCAGGGGAATAATGCCTGTTGGTGGTCAGCATCCTCCTAGCTCG

GGGGTTGGGGCTAGGCTTGAGTGGTAGAAGGCCCAGAAGAAGCCTGCAAA

AAAGAATACTTCTGAGATGATAAAGAGGATTATTCCGTATCAAAGGCCTT

TTTGAACGATAGGTGTATGATGGCCTTGAAATGGGCTTTCTCGGATA

>FC4_Fd_mt_F

TAACACTATTTATTGTATTCCAACTAAAAATCTCAAAATACTTATATCCA

TCAAACCCAGAACCTAAATCCATAACCACACTAAAACAACTTAATCCCTG

AGAAAAAAAATGAACGAAAATCTATTCGCCTCTTTCACTACCCCAACAAT

AATAGGATTACCTATTGTTATTTTAATTATTATATTTCCAAGTATTTTAT

TCCCTTCACCTAACCGACTAATTAATAACCGTCTAGTTTCACTCCAACAA

TGACTAGTACAATTAACATCAAAACAAATACTGGCTATTCATAATCATAA

AGGACAAACCTGAGCCCTAATACTAATGTCCCTAATTCTATTTATTGGGT

CAACAAACTTATTAGGCCTATTACCCCACTCATTCACCCCAACTACCCAA

CTATCAATAAATTTAGGAATAGCTATCCCGCTATGAGCCGGCACTGTAAT

TACCGGGTTTCGCCACAAGACTAAAGCATCTCTGGCCCACTTTCTACCAC

AAGGAACACCTGTCCCCCTAATTCCTATGCTTGTAGTCATTGAGACTATC

AGCCTCTTTATCCAACCCATAGCTCTCGCCGTACGACTTACAGCCAACAT

CACCGCAGGTCACTTATTAATACATCTAATTGGAGGAGCCGCCCTAGCCC

TGATAAACATTAGCACCTCTATTGCCTTAATCACCTTTACCATTCTCATT

TTACTAACAATCCTTGAATTTGCCGTAGCCCTAATCCAAGCCTATGTTTT

TACCCTGCTAGTAAGCCTATACTTACATGATAACACCTAATGACCCACCA

AACCCATGCATACCACATAGTCAACCCTAGCCCATGGCCACTTACAGGAG

CCCTTTCAGCCCTCTTAATAACCTCAGGCCTGGCTATATGATTCCACTAC

AACTCAACACTGCTGTTAACCCTTGGAATAACTACCACCTTACT

>FC4_Fd_mt_R

ATGATGTTTCATAGTATTCGGAGGCTTGAAGGAGTGTAAAGTAGACCCCT

AAAGAGATTGTAATAAATAGTGCTTGAAGCATGTGTTTTCGATTTCCCTC

CATCAAACTGTGGTGAGCTCAGGTGATTGATACTCCGGAGGCTAGAAGCA

CGGAGGTATTAAGTAGTGGAACTTCCAGGGGGTTCAGGGGAATAATGCCT

GTTGGTGGTCAGCATCCTCCTAGCTCGGGGGTTGGGGCTAAGCTTGAGTG G

>FC12_Fd_mt_F

AACACTATTTATTGTATTCCAACTAAAAATCTCAAAATACTTATATCCAT

CAAACCCAGAACCTAAATCCATAACCACACTAAAACAACTTAATCCCTGA

GAAAAAAAATGAACGAAAATCTATTCGCCTCTTTCACTACCCCAACAATA

ATAGGATTACCTATTGTTATTTTAATTATTATATTTCCAAGTATTTTATT

CCCTTCACCTAACCGACTAATTAATAACCGTCTAGTTTCACTCCAACAAT

GACTAGTACAATTAACATCAAAACAAATACTGGCTATTCATAATCATAAA

GGACAAACCTGAGCCCTAATACTAATGTCCCTAATTCTATTTATTGGGTC

AACAAACTTATTAGGCCTATTACCCCACTCATTCACCCCAACTACCCAAC

TATCAATAAATTTAGGAATAGCTATCCCGCTATGAGCCGGCACTGTAATT

ACCGGGTTTCGCCACAAGACTAAAGCATCTCTGGCCCACTTTCTACCACA

GGGAACACCTGTCCCCCTAATTCCTATGCTTGTAGTCATTGAGACTATCA

GCCTCTTTATCCAACCCATAGCTCTCGCCGTACGACTTACAGCCAACATC

ACCGCAGGTCACTTATTAATACATCTAATTGGAGGAGCCGCCCTAGCCCT

GATAAACATTAGCACCTCTATTGCCTTAATCACCTTTACCATTCTCATTT

TACTAACAATCCTTGAATTTGCCGTAGCCCTAATCCAAGCCTATGTTTTT

ACCCTGCTAGTAAGCCTATACTTACATGATAACACCTAATGACCCACCAA

ACCCATGCATACCACATAGTCAACCCTAGCCCATGGCCACTTACAGGAGC

CCTTTCAGCCCTCTTAATAACCTCAGNCCTGGCTATATGATTCCACTACA

ACTCAACACTGCTGTTAACCCTTGGAATAACTACCAACCTACTAACTATA

TATCAATGATGACGA

>FC12_Fd_mt_R

ATGATGTTTCATAGTATTCGGAGGCTTGAAGGAGTGTAAAGTAGACTCCT

AANNAGATTGTAATAAATAGTGCTTGAAGCATGTGTTTTCGATTTCCCTC

CATCAAACTGTGGTGAGCTCAGGTGATTGATACTCCGGAGGCTAGAAGCA

CGGAGGTATTAAGTAGTGGAACTTCCAGGGGGTTCAGGGGAATAATGCCT

GTTGGTGGTCAGCATCCTCCTAGCTCGGGGGTTGGGGCTAGGCTTGAGTG

GTAGAAGGCCCAGAAGAAGCCTGCGAAAAAGAATACTTCTGAGATGATAA

AGAGGATTATTCCGTATCGAAGGCCTTTTTGAACGATAGGTGTATGATGG

CCTTGGAATGTGCTTTCTCGGATAATGTCTCGTCATCATTGATATATAGT

TAGTAGGTTGGTAGTTATTCCAAGGGTTAACAGCAGTGTTGAGTTGTAGT

GGAATCATATAGCCAGGCCTGAGGTTATTAAGAGGGCTGAAAGGGCTCCT

GTAAGTGGCCATGGGCTAGGGTTGACTATGTGGTATGCATGGGTTTGGTG

GGTCATTAGGTGTTATCATGTAAGTATAGGCTTACTAGCAGGGTAAAAAC

ATAGGCTTGGATTAGGGCTACGGCAAATTCAAGGATTGTTAGTAAAATGA

GAATGGTAAAGGTGATTAAGGCAATAGAGGTGCTAATG

>FC13_Fd_mt_F

AATCTCAAAATACTTATATCCATCAAACCCAGAACCTAAATCCATAACCA

CACTAAAACAACTTAATCCCTGAGAAAAAAAATGAACGAAAATCTATTCG

CCTCTTTCACTACCCCAACAATAATAGGATTACCTATTGTTATTTTAATT

ATTATATTTCCAAGTATTTTATTCCCTTCACCTAACCGACTAATTAATAA

CCGTCTAGTTTCACTCCAACAATGACTAGTACAATTAACATCAAAACAAA

TACTGGCTATTCATAATCATAAAGGACAAACCTGAGCCCTAATACTAATG

TCCCTAATTCTATTTATTGGGTCAACAAACTTATTAGGCCTATTACCCCA

CTCATTCACCCCAACTACCCAACTATCAATAAATTTAGGAATAGCTATCC

CGCTATGAGCCGGCACTGTAATTACCGGGTTTCGCCACAAGACTAAAGCA

TCTCTGGCCCACTTTCTACCACAGGGAACACCTGTCCCCCTAATTCCTAT

GCTTGTAGTCATTGAGACTATCAGCCTCTTTATCCAACCCATAGCTCTCG

CCGTACGACTTACAGCCAACATCACCGCAGGTCACTTATTAATACATCTA

ATTGGAGGAGCCGCCCTAGCCCTGATAAACATTAGCACCTCTATTGCCTT

AATCACCTTTACCATTCTCATTTTACTAACAATCCTTGAATTTGCCGTAG

CCCTAATCCAAGCCTATGTTTTTACCCTGCTAGTAAGCCTATACTTACAT

GATAACACCTAATGACCCACCAAACCCATGCATACCACATAGTCAACCCT

AGCCCATGGCCACTTACAGGAGCCCTTTCAGCCCTTTTAATAACCTCAGG

CCTGGCTATATGATTCCACTACAACTCAACACTGCTGTTAACCCTTGGAA

TAACTACCAACCTACTAACTATATATCAATGATGACGANACATTA

>FC13_Fd_mt_R

GAAGGAGTGTAAAGTAGACTCCTAAAGAGATTGTAATAAATAGTGCTTGA

AGCATGTGTTTTCGATTTCCCTCCATCAAACTGTGGTGAGCTCAGGTGAT

TGATACTCCGGAGGCTAGAAGCACGGAGGTATTAAGTAGTGGAACTTCCA

GGGGGTTCAGGGGAATAATGCCTGTTGGTGGTCAGCATCCTCCTAGCTCG

GGGGTTGGGGCTAGGCTTGAGTGGTAGAAGGCCCAGAAGAAGCCTGCGAA

AAAGAATACTTCTGAGATGATAAAGAGGATTATTCCGTATCGAAGGCCTT

TTTGAACGATAGGTGTATGATGGCCTTGGAATGTGCTTTCTCGGATAATG

TCTCGTCATCATTGATATATAGTTAGTAGGTTGGTAGTTATTCCAAGGGT

TAACAGCAGTGTTGAGTTGTAGTGGAATCATATAGCCAGGCCTGAGGTTA

TTAAAAGGGCTGAAAGGGCTCCTGTAAGTGGCCATGGGCTAGGGTTGACT

ATGT

>FC15_Fd_mt_F

TCCAACTAAAAATCTCAAAATACTTATATCCGTCAAACCCAGAACCTAAA

TCCATAACCACACTAAAACAACTTAATCCCTGAGAAAAAAAATGAACGAA

AATCTATTCGCCTCTTTCACTACCCAAACAATAATAGGATTACCTATTGT

TATTTTAATTATTATATTTCCAAGTATTTTATTCCCTTCACCTAACCGAC

TAATTAATAACCGTCTAGTTTCACTCCAACAATGACTAGTACAATTAACA

TCAAAACAAATACTGGCTATTCATAATCATAAAGGACAAACCTGAGCCCT

AATACTAATGTCCCTAATTCTATTTATTGGGTCAACAAACTTATTAGGCC

TATTACCCCACTCATTCACCCCAACTACCCAACTATCAATAAATTTAGGA

ATAGCTATCCCGCTATGAGCCGGCACTGTAATTACCGGGTTTCGCCACAA

GACTAAAGCATCTCTGGCCCACTTTCTACCACAAGGAACACCTGTCCCCC

TAATTCCTATGCTTGTAGTCATTGAGACTATCAGCCTCTTTATCCAACCC

ATAGCTCTCGCCGTACGACTTACAGCCAACATCACCGCAGGTCACTTATT

AATACATCTAATTGGAGGAGCCGCCCTAGCCCTGATAAACA

>FC15_Fd_mt_R

GATGATGTTTCATAGTATTCGGAGGCTTGAAGGAGTGTAAAGTAGACCCC

TAAAGAGATTGTAATAAATAGTGCTTGAAGCATGTGTTTTCGATTTCCCT

CCATCAAACTGTGGTGAGCTCAGGTGATTGATACTCCGGAGGCTAGAAGC

ACGGAGGTATTAAGTAGTGGAACTTCCAGGGGGTTCAGGGGAATAATGCC

TGTTGGTGGTCAGCATCCTCCTAGCTCGGGGGTTGGGGCTAGGCTTGAGT

GGTAGAAGGCCCAGAAGAAGCCTGCGAAAAAGAATACTTCTGAGATGATA

AAGAGGATTATTCCGTATCGAAGGCCTTTTTGAACGATAGGTGTATGATG

GCCTTGGAATGTGCTTTCTCGGATAATGTCTC

>FC16_Fd_mt_F

ATAACACTATTTATTGTATTCCAACTAAAAATCTCAAAATACTTATATCC

ATCAAACCCAGAACCTAAATCCATAACCACACTAAAACAACTTAATCCCT

GAGAAAAAAAATGAACGAAAATCTATTCGCCTCTTTCACTACCCCAACAA

TAATAGGATTACCTATTGTTATTTTAATTATTATATTTCCAAGTATTTTA

TTCCCTTCACCTAACCGACTAATTAATAACCGTCTAGTTTCACTCCAACA

ATGACTAGTACAATTAACATCAAAACAAATACTGGCTATTCATAATCATA

AAGGACAAACCTGAGCCCTAATACTAATGTCCCTAATTCTATTTATTGGG

TCAACAAACTTATTAGGCCTATTACCCCACTCATTCACCCCAACTACCCA

ACTATCAATAAATTTAGGAATAGCTATCCCGCTATGAGCCGGCACTGTAA

TTACCGGGTTTCGCCACAAGACTAAAGCATCTCTGGCCCACTTTCTACCA

CAGGGAACACCTGTCCCCCTAATTCCTATGCTTGTAGTCATTGAGACTAT

CAGCCTCTTTATCCAACCCATAGCTCTCGCCGTACGACTTACAGCCAACA

TCACCGCAGGTCACTTATTAATACATCTAATTGGAGGAGCCGCCCTAGCC

CTGATAAACATTAGCACCTCTATTGCCTTAATCACCTTTACCATTCTCAT

TTTACTAACAATCCTTGAATTTGCCGTAGCCCTAATCCAAGCCTATGTTT

TTACCCTGCTAGTAAGCCTATACTTACATGATAACACCTAATGACCCACC

AAACCCATGCATACCACATAGTCAACCCTAGCCCATGGCCACTTACAGGA

GCCCTTTCAGCCCTCTTAATAACCTCAGGCCTGGCTATATGATTCCACTA

CAACTCAACACTGCTGTTAACCCTTGGAATAACTACCAACCTACTAACTA

TATATCAATGATGACGAGACATTATCCGAGAAAGCACATTCC

>FC16_Fd_mt_R

ATGATGTTTCATAGTATTCGGAGGCTTGAAGGAGTGTAAAGTAGACTCCT

AAAGAGATTGTAATAAATAGTGCTTGAAGCATGTGTTTTCGATTTCCCTC

CATCAAACTGTGGTGAGCTCAGGTGATTGATACTCCGGAGGCTAGAAGCA

CGGAGGTATTAAGTAGTGGAACTTCCAGGGGGTTCAGGGGAATAATGCCT

GTTGGTGGTCAGCATCCTCCTAGCTCGGGGGTTGGGGCTAGGCTTGAGTG

GTAGAAGGCCCAGAAGAAGCCTGCGAAAAAGAATACTTCTGAGATGATAA

AGAGGATTATTCCGTATCGAAGGCCTTTTTGAACGATAGGTGTATGATGG

CCTTGGAATGTGCTTTCTCGGATAATGTCTCGTCATCATTGATATATAGT

TAGTAGGTTGGTAGTTATTCCAAGGGTTAACAGCAGTGTTGAGTTGTAGT

GGAATCATATAG

>FC17_Fd_mt_F

CTATTTATTGTAGTCCAACTAAAAATCTCAAAATACTTATATCCATCAAA

CCCAGAACCTAAATCCATAACCACACTAAAACAACTTAATCCCTGAGAAA

AAAAATGAACGAAAATCTATTCGCCTCTTTCACTACCCCAACAATAATAG

GATTACCTATTGTTATTTTAATTATTATATTTCCAAGTATTTTATTCCCT

TCACCTAACCGACTAATTAATAACCGTCTAGTTTCACTCCAACAATGACT

AGTACAATTAACATCAAAACAAATACTGGCTATTCATAATCATAAAGGAC

AAACCTGAGCCCTAATACTAATGTCCCTAATTCTATTTATTGGGTCAACA

AACTTATTAGGCCTATTACCCCACTCATTCACCCCAACTACCCAACTATC

AATAAATTTAGGAATAGCTATCCCGCTATGAGCCGGCACTGTAATTACCG

GGTTTCGCCACAAGACTAAAGCATCTCTGGCCCACTTTCTACCACAAGGA

ACACCTGTCCCCCTAATTCCTATGCTTGTAGTCATTGAGACTATCAGCCT

CTTTATCCAACCCATAGCTCTCGCCGTACGACTTACAGCCAACATCACCG

CAGGTCACTTATTAATACATCTAATTGGAGGAGCCGCCCTAGCCCTGATA

AACATTAGCACCTCTATTGCCTTAATCACCTTTACCATTCTCATTTTACT

AACAATCCTTGAATTTGCCGTAGCCCTAATCCAAGCCTATGTTTTTACCC

TGCTAGTAAGCCTATACTTACATGATAACACCTAATGACCCACCAAACCC

ATGCATACCACATAGTCAACCCTAGCCCATGGCCACTTACAGGAGCCCTT

TCAGCCCTCTTAATAACCTCAGGCCTGGCTATATGATTCCACTACAACTC

AACACTGCTGTTAACCCTTGG

>FC17_Fd_mt_R

CATAGTATTCGGAGGCTTGAAGGAGTGTAAAGTAGACCCCTAAAGAGATT

GTAATAAATAGTGCTTGAAGCATGTGTTTTCGATTTCCCTCCATCAAACT

GTGGTGAGCTCAGGTGATTGATACTCCGGAGGCTAGAAGCACGGAGGTAT

TAAGTAGTGGAACTTCCAGGGGGTTCAGGGGAATAATGCCTGTTGGTGGT

CAGCATCCTCCTAGCTCGGGGGTTGGGGCTAGGCTTGAGTGGTAGAAGGC

CCAGAAGAAGCCTGCGAAAAAGAATACTTCTGAGATGATAAAGAGGATTA

TTCCGTATCGAAGGCCTTTTTGAACGATAGGTGTATGATGGCCTTGGAAT

GTGCTTTCTCGGATAATGTCTCGTCATCATTGATATATAGTTAGTAGGTT

GGTAGTTATTCCAAGGGTTAACAGCAGTGTTGAGTTGTAGTGGAATCATA

TAGCCAGGCCTGAGGTTATTAAGAGGGCTGAAAGGGCTCCTGTAAGTGGC

CATGGGCTAGGGTTGACTATGTGGTATGCATGGGTTTGGTGGGTCATTAG

GTGTTATCATGTAAGTATAGGCTTACTAG

>FC18_Fd_mt_F

CTCAAAATACTTATATCCATCAAACCCAGAACCTAAATCCATAACCACAC

TAAAACAACTTAATCCCTGAGAAAAAAAATGAACGAAAATCTATTCGCCT

CTTTCACTACCCCAACAATAATAGGATTACCTATTGTTATTTTAATTATT

ATATTTCCAAGTATTTTATTCCCTTCACCTAACCGACTAATTAATAACCG

TCTAGTTTCACTCCAACAATGACTAGTACAATTAACATCAAAACAAATAC

TGGCTATTCATAATCATAAAGGACAAACCTGAGCCCTAATACTAATGTCC

CTAATTCTATTTATTGGGTCAACAAACTTATTAGGCCTATTACCCCACTC

ATTCACCCCAACTACCCAACTATCAATAAATTTAGGAATAGCTATCCCGC

TATGAGCCGGCACTGTAATTACCGGGTTTCGCCACAAGACTAAAGCATCT

CTGGCCCACTTTCTACCACAGGGAACACCTGTCCCCCTAATTCCTATGCT

TGTAGTCATTGAGACTATCAGCCTCTTTATCCAACCCATAGCTCTCGCCG

TACGACTTACAGCCAACATCACCGCAGGTCACTTATTAATACATCTAATT

GGAGGAGCCGCCCTAGCCCTGATAAACATTAGCACCTCTATTGCCTTAAT

CACCTTTACCATTCTCATTTTACTAACAATCCTTGAATTTGCCGTAGCCC

TAATCCAAGCCTATGTTTTTACCCTGCTAGTAAGCCTATACTTACATGAT

AACACCTAATGACCCACCAAACCCATGCATACCACATAGTCAACCCTAGC

CCATGGCCACTTACAGGAGCCCTTTCAGCCCTCTTAATAACCTCAGGCCT

GGCTATATGATTCCACTACAACTCAACACTGCTGTTAACCCTTGGA

>FC18_Fd_mt_R

GGAGTGTAAAGTAGACTCCTAAAGAGATTGTAATAAATAGTGCTTGAAGC

ATGTGTTTTCGATTTCCCTCCATCAAACTGTGGTGAGCTCAGGTGATTGA

TACTCCGGAGGCTAGAAGCACGGAGGTATTAAGTAGTGGAACTTCCAGGG

GGTTCAGGGGAATAATGCCTGTTGGTGGTCAGCATCCTCCTAGCTCGGGG

GTTGGGGCTAGGCTTGAGTGGTAGAAGGCCCAGAAGAAGCCTGCGAAAAA

GAATACTTCTGAGATGATAAAGAGGATTATTCCGTATCGAAGGCCTTTTT

GAACGATAGGTGTATGATGGCCTTGGAATGTGCTTTCTCGGATAATGTCT

CGTCATCATTGATATATAGTTAGTAGGTTGGTAGTTATTCCAAGGGTTAA

CAGCAGTGTTGAGTTGTAGTGGAATCATATAGCCAGGCCTGAGGTTATTA

AGAGGGCTGAAAGGGCTCCTGTAAGTGGCCATGGGCTAGGGTTGACTATG

TGGTATGCATGGGTTTGGTGGGTCATTAGGTGTTATCATGTAA

>FC19_Fd_mt_F

TCCAACTGAAAATCTCAAAATACTTATATCCATCAAACCCAGAACCTAAA

TCCATAACCACACTAAAACAACTTAATCCCTGAGAAAAAAAATGAACGAA

AATCTATTCGCCTCTTTCACTACCCCAACAATAATAGGATTACCTATTGT

TATTTTAATTATTATATTTCCAAGTATTCTATTCCCTTCACCTAACCGAC

TAATTAATAACCGTCTAGTTTCACTCCAACAATGACTAGTACAATTAACA

TCAAAACAAATACTGGCTATTCATAATCATAAAGGACAAACCTGAGCCCT

AATACTAATGTCCCTAATTCTATTTATTGGGTCAACAAACTTATTAGGCC

TATTACCCCACTCATTCACCCCAACTACCCAATTATCAATAAATTTAGGA

ATAGCTATCCCGCTATGAGCCGGCACTGTAATTACCGGGTTTCGCCACAA

GACTAAAGCATCTCTGGCCCACTTTCTACCACAAGGAACACCTGTCCCCC

TAATTCCTATGCTTGTAGTCATTGAGACTATCAGCCTCTTTATCCAACCC

ATAGCTCTCGCCGTACGACTTACAGCCAACATCACCGCAGGTCACTTATT

AATACATCTAATTGGAGGAGCCGCCCTAGCCCTGATAAATATTAGCACCT

CTATTGCCTTAATCACCTTTACCATTCTCATTTTACTAACAATCCTTGAA

TTTGCCGTAGCCCTAATCCAAGCCTATGTTTTTACCCTGCTAGTAAGCCT

ATACTTACATGATAACACCTAATGACCCACCAAACCCATGCATACCACAT

AGTCAACCCTAGCCCATGGCCACTTAC

>FC19_Fd_mt_R

GGAGGAGTGTAAAGTAGACCCCTAAAGAGATTGTAATAAATAGTGCTTGG

AGCATGTGTTTTCGATTTCCTTCCATCAAACTGTGGTGAGCTCAGGTGAT

TGATACTCCGGAGGCTAGAAGCACGGAGGTATTAAGTAGTGGAACTTCCA

GGGGGTTCAGGGGAATAATGCCTGTCGGTGGTCAGCATCCTCCTAGCTCG

GGGGTTGGGGCTAGGCTTGAGTGGTAGAAGGCCCAGAAGAAGCCTGCGAA

AAAGAATACTTCTGAGATGATAAAGAGGATTATTCCGTATCGAAGGCCTT

TTTGAACGATAGGTGTATGATGGCCTTGGAATGTGCTTTCTCGGATAATG

TCTCGTCATCATTGATATATAGTTAGTAGGTTGGTAGTTATTCCAAGGGT

TAACAGCAGTGTTGAGTTGTAGTGGAATCATATAGCCAGGCCTGAGGTTA

TTAAGAGGGCTGAAAGGGCTCCTGTAAGTGGCCATGGGCTAGGGTTGACT

ATGTGGTATGCATGGGTTTGGTGGGTCATTAGGTGTTATCATGTAAGTAT

AGGCTTACTAGCAGGGTAAAAACATAGGCTTGGATTAGGG

>FC20_Fd_mt_F

CCCACTAATAATCTCAAAATACTTATATCCATCAAACCCAGAACCTAAAT

CCATAACCACACTAAAACAACTTAATCCCTGAGAAAAAAAATGAACGAAA

ATCTATTCGCCTCTTTCACTACCCCAACAATAATAGGATTACCTATTGTT

ATTTTAATTATTATATTTCCAAGTATTTTATTCCCTTCACCTAACCGACT

AATTAATAACCGTCTAGTTTCACTCCAACAATGACTAGTACAATTAACAT

CAAAACAAATACTGGCTATTCATAATCATAAAGGACAAACCTGAGCCCTA

ATACTAATGTCCCTAATTCTATTTATTGGGTCAACAAACTTATTAGGCCT

ATTACCCCACTCATTCACCCCAACTACCCAACTATCAATAAATTTAGGAA

TAGCTATCCCGCTATGAGCCGGCACTGTAATTACCGGGTTTCGCCACAAG

ACTAAAGCATCTCTGGCCCACTTTCTACCACAGGGAACACCTGTCCCCCT

AATTCCTATGCTTGTAGTCATTGAGACTATCAGCCTCTTTATCCAACCCA

TAGCTCTCGCCGTACGACTTACAGCCAACATCACCGCAGGTCACTTATTA

ATACATCTAATTGGAGGAGCCGCCCTAGCCCTGATAAACATTAGCACCTC

TATTGCCTTAATCACCTTTACCATTCTCATTTTACTAACAATCCTTGAAT

TTGCCGTAGCCCTAATCCAAGCCTATGTTTTTACCCTGCTAGTAAGCCTA

TACTTACATGATAACACCTAATGACCCACCAAACCCATGCATACCACATA

GTCAACCCTAGCCCATGGCCACTTACAGGAGCCCTTTCAGCCCTCTTAAT

AACCTCAGGCCTGGCTATATGATTCCACTACAACTCAACACTGCTGTTAA

CCCTTGGAATAACTACCAACCTACTAACTAT

>FC20_Fd_mt_R

AAGGAGTGTAAAGTAGACTCCTAAAGAGATTGTAATAAATAGTGCTTGAA

GCATGTGTTTTCGATTTCCCTCCATCAAACTGTGGTGAGCTCAGGTGATT

GATACTCCGGAGGCTAGAAGCACGGAGGTATTAAGTAGTGGAACTTCCAG

GGGGTTCAGGGGAATAATGCCTGTTGGTGGTCAGCATCCTCCTAGCTCGG

GGGTTGGGGCTAGGCTTGAGTGGTAGAAGGCCCAGAAGAAGCCTGCGAAA

AAGAATACTTCTGAGATGATAAAGAGGATTATTCCGTATCGAAGGCCTTT

TTGAACGATAGGTGTATGATGGCCTTGGAATGTGCTTTCTCGGATAATGT

CTCGTCATCATTGATATATAGTTAGTAGGTTGGTAGTTATTCCAAGGGTT

AACAGCAGTGTTGAGTTGTAGTGGAATCATATAGCCAGGCCTGAGGTTAT

TAAGAGGGCTGAAAGGGCTCCTGTAAGTGGCCATGGGCTAGGGTTGACTA

TGTGGTATGCATGGGTTTGGTGGGTCATTAGGTGTTATCATGTAAGTATA

GGCTTACTAGCAGGGTAAAAACATAGGCTTGGATTAGGGCTACGGCAAAT

TCAAGGATTGTTAGTAAAATGAGAATGGTAAAGGTGATTAAGGCAATAGA GGT

>FC22_Fd_mt_F

CTATTTATTGTATTCCAACTAAAAATCTCAAAATACTTATATCCATCAAA

CCCAGAACCTAAATCCATAACCACACTAAAACAACTTAATCCCTGAGAAA

AAAAATGAACGAAAATCTATTCGCCTCTTTCACTACCCCAACAATAATAG

GATTACCTATTGTTATTTTAATTATTATATTTCCAAGTATTTTATTCCCT

TCACCTAACCGACTAATTAATAACCGTCTAGTTTCACTCCAACAATGACT

AGTACAATTAACATCAAAACAAATACTGGCTATTCATAATCATAAAGGAC

AAACCTGAGCCCTAATACTAATGTCCCTAATTCTATTTATTGGGTCAACA

AACTTATTAGGCCTATTACCCCACTCATTCACCCCAACTACCCAACTATC

AATAAATTTAGGAATAGCTATCCCGCTATGAGCCGGCACTGTAATTACCG

GGTTTCGCCACAAGACTAAAGCATCTCTGGCCCACTTTCTACCACAAGGA

ACACCTGTCCCCCTAATTCCTATGCTTGTAGTCATTGAGACTATCAGCCT

CTTTATCCAACCCATAGCTCTCGCCGTACGACTTACAGCCAACATCACCG

CAGGTCACTTATTAATACATCTAATTGGAGGAGCCGCCCTAGCCCTGATA

AACATTAGCACCTCTATTGCCTTAATCACCTTTACCATTCTCATTTTACT

AACAATCCTTGAATTTGCCGTAGCCCTAATCCAAGCCTATGTTTTTACCC

TGCTAGTAAGCCTATACTTACATGATAACACCTAATGACCCACCAAACCC

ATGCATACCACATAGTCAACCCTAGCCCATGGCCACTTACAGGAGCCCTT

TCAGCCCTCTTAATAACCTCAGGCCTGGCTATATGATTCCACTACAACTC

AACACTGCTGTTAACCCTTGGAATAACTACCAACCTACTAACTATATATC AATGATG

>FC22_Fd_mt_R

GATGATGTTTCATAGTATTCGGAGGCTTGAAGGAGTGTAAAGTAGACCCC

TAAAGAGATTGTAATAAATAGTGCTTGAAGCATGTGTTTTCGATTTCCCT

CCATCAAACTGTGGTGAGCTCAGGTGATTGATACTCCGGAGGCTAGAAGC

ACGGAGGTATTAAGTAGTGGAACTTCCAGGGGGTTCAGGGGAATAATGCC

TGTTGGTGGTCAGCATCCTCCTAGCTCGGGGGTTGGGGCTAGGCTTGAGT

GGTAGAAGGCCCAGAAGAAGCCTGCGAAAAAGAATACTTCTGAGATGATA

AAGAGGATTATTCCGTATCGAAGGCCTTTTTGAACGATAGGTGTATGATG

GCCTTGGAATGTGCTTTCTCGGATAATGTCTCGTCATCATTGATATATAG

TTAGTAGGTTGGTAGTTATTCCAAGGGTTAACAGCAGTGTTGAGTTGTAG

TGGAATCATATAGCCAGGCCTGAGGTTATTAAGAGGGCTGAAAGGGCTCC

TGTAAGTGGCCATGGGCTAGGGTTGACTATGTGGTA